

巡回群行列を用いた回転物体画像の複素部分空間への投影とその応用

玉木 徹[†] 天野 敏之^{††} 金田 和文[†]

[†] 広島大学大学院工学研究科情報工学専攻 〒739-8527 広島県東広島市鏡山 1-4-1

^{††} 奈良先端科学技術大学院大学情報科学研究科 〒630-0101 奈良県生駒市高山町 8916 番地の 5

E-mail: [†]{tamaki,kin}@hiroshima-u.ac.jp, ^{††}amano@is.naist.jp

あらまし 本論文では、画像間の補間や姿勢推定に有効であろう線形写像の一つとして、1 軸回転する物体のような巡回する画像列が与えられたときに、画像間の関係を単純な 1 次元複素空間での回転で表現できるような部分空間への線形写像を提案する。提案手法は、まず画像列を置換する巡回群を考え、その群を表現する行列を低ランクの複素行列の積に分解する。この複素行列により、画像空間中の画像が複素部分空間へと投影されるが、その部分空間においては画像間の関係が単純な複素平面上での回転で表現されることを示す。そしてサンプル画像間を補間する方法と実画像を用いた例を示す。

キーワード 姿勢推定, 巡回群, 回帰, 線形写像, 複素対角化, 複素部分空間, パラメトリック固有空間法, 画像生成

Projecting images of rotating object onto complex subspace by the matrix of cyclic group and its application

Toru TAMAKI[†], Toshiyuki AMANO^{††}, and Kazufumi KANEDA[†]

[†] Department of Information Engineering, Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-hiroshima, Hiroshima, 739-8527 Japan

^{††} Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: [†]{tamaki,kin}@hiroshima-u.ac.jp, ^{††}amano@is.naist.jp

Abstract This paper proposes a linear map that projects images of an rotating object about an axis onto a complex subspace where the rotation is represented with an simple rotation in one dimensional complex space. First we consider a cyclic group that permutes the image sequence, then construct a complex matrix of the cyclic group by multiplication of low rank matrices. We demonstrate interpolation of the image sequence by rotation in the subspace, and discuss the ability of the proposed linear map onto the complex subspace.

Key words pose estimation, cyclic group, regression, linear map, complex diagonalization, complex subspace, parametric eigenspace, view generation

1. はじめに

画像に写る物体の姿勢パラメータ推定はコンピュータビジョンの重要な問題として多くの研究がなされている。その多くは既知形状の剛体を仮定し、運動や投影の解析的モデルを採用して最適化問題を解く [1] ものである。それに対して村瀬ら [2], [3] のパラメトリック固有空間法は、物体の形状や 3 次元空間での投影といった幾何学的な撮像過程を仮定せず、連続的に変化するパラメータ列とそれに付随する画像系列を学習セットとし、その画像列が固有空間中に描く軌跡 (多様体) を認識に用いる、いわばパターン認識的姿勢推定問題を提起した。

多様体を描く非線形性の強い画像列のようなデータを教師なしのデータ解析問題として扱うために、カーネルトリックを利用した非線形主成分分析 (kernel PCA, kPCA) [4], [5] や SVM [6] などの手法や、多様体学習 (manifold learning) [7] が提案されてきている。これらは、固有顔 [8], [9] に代表されるような、主成分分析や判別分析を応用した線形的な認識手法よりも優れた結果を出すことが示されている。

一方、パラメトリック固有空間法のような姿勢推定手法が扱うのは、ある画像とそのときの姿勢パラメータというセットを学習する教師付き問題であり、すなわちデータのパラメータへの回帰問題に相当する。非線形な回帰手法にはカーネルトリッ

クを用いた SVM 回帰 [10] やカーネルリッジ回帰 [11] があり、それらを用いた物体の姿勢推定 [12] が提案されている。また、線形的な手法 [13], [14] も提案されている。

認識問題については非線形手法の優位性が示されているが、回帰問題についてはあまり議論がされていない。カーネルを用いた回帰自体の性能評価 [11] については議論されているものの、実際の画像の姿勢推定の状況における評価はなされていない。

しかし、画像の姿勢推定問題において本当に非線形手法が必要であろうか。一般の回帰問題と異なり、姿勢推定問題では画像の次元（画素数）が画像の枚数（サンプル数）よりもかなり大きい。つまり、高次元空間中にかなり疎らにサンプル点が存在することになる。そのため、物体の画像が多様体を描いたとしても、膨大なサンプルを集めない限り、実際には高次元空間中の点と点を結び折れ線でしかなく、多様体としての振る舞いを眺めることはできない。

画像の次元がサンプル数よりも大きいという状況においては、単にサンプルが部分空間に収まってしまおうというだけではない。この状況では、線形手法 [13], [14] における推定のための連立方程式の解は、最小ノルム解、であることがすでに指摘されている [15], [16]。つまり、学習サンプルにおいては厳密に方程式が成り立つ解が無数に存在し、その中のノルムが最小である解を 1 つ推定しているに過ぎない。したがって、1 軸回転する物体の画像列の姿勢パラメータを推定するという問題は、画像列を 2 次元平面の単位円周上に等間隔に並ぶ点への投影を実現する（無数の写像のうちの一つの）線形写像を求める問題とみなすことができる [16]。もしサンプル数が画像の次元よりも多い場合でも、非線形に次元を上げてしまえばその条件を満たすため、同様に単なる線形写像の推定になる。

そのため、おそらく姿勢推定に関しては、複雑な非線形手法までは必要ではなく、線形手法による推定でも十分であるといえるであろう、と思われる。ただし汎化性能には違いがあることが予想されるためその評価は必要ではあるが、前述の通りあまり議論されていないのが現状である。

そこで、非線形手法は議論せずに線形手法に焦点をあて、姿勢推定などに有効な画像の線形写像を追求しようというのが本論文の狙いである。どのような線形手法や線形写像が有効なのかという議論は十分になされていないにもかかわらず、線形手法の改良という場合には単にカーネル法を適用して手法を非線形にしてしまうことが多い。しかし画像の次元がサンプル数よりも大きい場合には、検討すべき線形写像が多数存在するはずである。それらを検討した後、有効な線形手法を非線形に拡張することで、画像の姿勢推定という特有の問題に対する手法に拡張できるとと思われる。

本論文では、画像間の補間や姿勢推定に有効であろう線形写像の一つとして、1 軸回転する物体のような巡回する画像列が与えられたときに、画像間の関係を単純な 1 次元複素空間での回転で表現できるような部分空間への線形写像を提案する。提案手法は、まず画像列を置換する巡回群を考え、その群を表現する行列を低ランクの複素行列の積に分解する。この複素行列により、画像空間中の画像が複素部分空間へと投影されるが、

その部分空間においては画像間の関係が単純な複素平面上での回転で表現されることを示す。

本論文の構成は以下の通りである。まず 2. 節において、画像間の遷移を表現する巡回群を定義し、その群を表現する行列を導出する。次にその行列の低ランクの行列の積への分解について 3. 節で述べる。そして 4. 節において、サンプル画像間を補間する方法について説明し、4.2 節で実画像を用いた例を示す。

2. 画像列の巡回群による表現

ここでは、 n 枚の画像 x_1, x_2, \dots, x_n が与えられたとき、巡回群の行列表現によって画像間の関係を表すことを考える。まず巡回群について説明した後、どのように画像 x_1 と x_2 の間を補間して画像（例えば $x_{1.5}$ など）を生成するかについて述べ、それを応用してサンプル画像間を補間する手法について述べる。著者らはブロック対角化を用いた手法 [17], [18] を提案しているが、本稿では複素対角化を用いた手法について述べる。

2.1 巡回群による画像間の関係

画像 $x_j = (x_{j1}, x_{j2}, \dots, x_{jN})^T$ を N 次元ベクトルとし、その画像を表すパラメータを θ_j とする。1 軸回転を表す角度などのように、ここでは n 個のパラメータは連続的に $\theta_1, \theta_2, \dots, \theta_n$ と変化し、また θ_1 に循環して戻すものを考える。したがって、画像も x_1, x_2, \dots, x_n と変化し、また x_1 に戻るような順番付けがされているとする。さらに実際の状況として $N > n$ 、つまり画素数は画像数より大きいことを仮定する。

ここで、画像を x_1, x_2, \dots と変化させるために、以下のような群を考えよう。つまり、画像に左から作用する変換の群 $G_n = \{G, G^2, \dots, G^n\}$ があって、

$$x_2 = Gx_1 \quad (1)$$

$$x_3 = Gx_2 = GGx_1 = G^2x_1 \quad (2)$$

$$\vdots$$

$$x_n = Gx_{n-1} = \dots = G^{n-1}x_1 \quad (3)$$

$$x_1 = Gx_n = \dots = G^n x_1 \quad (4)$$

ここで、 G_n は n 位の有限巡回群、元 G は群 G_n の生成元と呼ばれる [19]。群の定義から、 G_n には各元の逆元が存在し、また単位元 I が存在する。

$$x_i = Gx_{i-1} = GG^{n-1}x_i = x_i, \quad G^{-1} = G^{n-1} \quad (5)$$

$$x_i = G^n x_i, \quad G^n = I \quad (6)$$

2.2 巡回群の行列表現

群自体は抽象的なものであるため、ベクトルとして扱われる画像 x_i を扱うために、群を行列で表現する。これは一般に群の表現論と呼ばれるものであるが、一般論には立ち入らずに、ここでは以下のように画像認識を容易にすることを目的に行列を構成することにする。

前節で述べた巡回群 G_n の生成元 G を行列 G で表現すると、画像 x_1, x_2, \dots の関係は以下のように表せる。

$$[x_2 \ x_3 \ \dots \ x_n \ x_1] = G[x_1 \ x_2 \ \dots \ x_{n-1} \ x_n] \quad (7)$$

ここで上式を

$$X_1 = GX_0 \quad (8)$$

とおく。 G の次元は $N \times N$ であるが、 $\text{rank}(X_1) = \text{rank}(X_0) = n$ であり、さらに $N > n$ を仮定しているので、連立方程式 (7) は劣決定問題になり、この方程式を満たす G は無数に存在する。そこで、ここでは最小ノルム型一般化逆行列 [20] を用いて以下のように G を定める。

$$G = X_1(X_0^T X_0)^{-T} X_0^T \quad (9)$$

したがって $\text{rank}(G) = \text{rank}(X_0) = n$ である。

2.3 行列の分解 ($N = n$ の場合)

一般的な状況では正方行列は対角化可能である [21] ので、本研究では G が対角化可能^(注1) であるとする。すると、 G は以下のように分解できる。

$$G = U^{-1}DU \quad (10)$$

$$D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N), \quad \lambda_{n+1} = \dots = \lambda_N = 0 \quad (11)$$

ここで D は対角行列で対角要素に固有値を持ち、一般に複素数である。また U も複素行列であるが、 G が対称行列ではないので、一般に U はユニタリ行列ではない。

ここで議論の簡単化のため、サンプル画像が十分にあって $N = n = \text{rank}(G)$ であるとしよう。すると、群 G_n の単位元 I は $N \times N$ 単位行列 I_N に相当するので、以下のように行列 G の冪乗は容易に計算できる。そこで、画像列を一周するために G の n 乗を考えると、以下ようになる。

$$G^n = I_N = U^{-1}D^nU \quad (12)$$

$$D^n = \text{diag}(\lambda_1^n, \lambda_2^n, \dots, \lambda_N^n) = UI_NU^{-1} = I_N \quad (13)$$

つまり、与えられた画像列を行列 G によって次々に遷移し、またもとの画像に戻るためには、 G の固有値は $\lambda_j^n = 1$ 、つまり 1 の原始 n 乗根でなければならない。

$$\lambda_j = \sqrt[n]{1} = e^{\frac{2k\pi}{n}i}, \quad i = \sqrt{-1}, \quad k = 1, 2, \dots, n \quad (14)$$

N 個の固有値 λ_j が、 n 通りあるうちのどの原始 n 乗根であるかは、元の与えられた画像列 X_0 に依存する。

2.4 群の作用としての複素回転

ここで再び G を式 (7) による画像 x_j への作用を見てみる。

$$x_{j+1} = Gx_j = U^{-1}DUx_j \quad (15)$$

$$Ux_{j+1} = DUx_j \quad (16)$$

$$x'_{j+1} = Dx'_j, \quad x'_j = Ux_j \quad (17)$$

$$\begin{pmatrix} x'_{j+1,1} \\ x'_{j+1,2} \\ \vdots \\ x'_{j+1,N} \end{pmatrix} = \begin{pmatrix} \lambda_1 x'_{j,1} \\ \lambda_2 x'_{j,2} \\ \vdots \\ \lambda_N x'_{j,N} \end{pmatrix} \quad (18)$$

(注1): 実数の範囲では 1×1 または 2×2 のブロック対角化 [17], [18] となるが、複素数の範囲では対角行列による対角化となる。

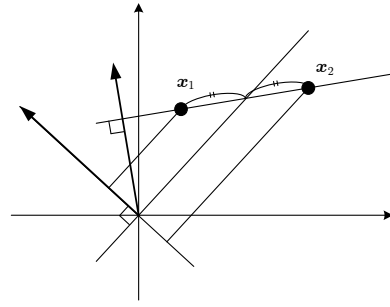


図1 2次元における複素回転の例

これは各画素を複素回転していることに他ならない。その解釈について以下に述べる。

複素数 z は、絶対値 $|z|$ と偏角 $\angle z$ を用いて $z = |z|e^{i\angle z}$ と表され、二つの複素数 z_1, z_2 の積は $z_1 z_2 = |z_1||z_2|e^{i(\angle z_1 + \angle z_2)}$ である。前節で述べたように、行列 G の固有値 λ_j は 1 の原始 n 乗根であるので、 $|\lambda_j| = 1$ である。よって、例えば式 (18) の 1 行目 $x'_{j+1,1} = \lambda_1 x'_{j,1}$ をみると、

$$|x'_{j+1,1}| = |x'_{j,1}|, \quad \angle x'_{j+1,1} = \angle x'_{j,1} + \frac{2\pi}{n}k \quad (19)$$

となる。つまり上の式は、固有値 λ_1 をかけることによって $x'_{j,1}$ が $x'_{j+1,1}$ になり、その絶対値は変わらず偏角だけが変わる (つまり複素平面における原点を中心とした回転) ことを意味している。

またここで $x'_1 = Ux_1$ であるので、 x'_1 は、実数である学習サンプル画像 $x_1 \in R^N$ を複素行列 $U \in C^{N \times N}$ により複素空間へと座標変換した複素画像 $x'_1 \in C^N$ であるとみなすことができる。

したがって複素行列 U は、単に各画素を複素回転させることで画像が x'_1, x'_2, \dots と次々に移りあうことができる複素空間への変換を実現している。

2.5 2次元における複素回転の例

前述のような複素行列による変換が存在することを、2次元の例を用いて示す。

$N = n = 2$ の場合、固有値は $\sqrt{1}$ 、つまり ± 1 である。したがって、2 サンプル x_1, x_2 を変換後した後の座標系においては、1 つ目の要素は同じ値を持ち ($x'_{11} = x'_{21}$)、2 つ目の要素は符号が逆転する ($x'_{12} = -x'_{22}$)。このような変換は簡単に示すことができる (図1 参照)。変換後の座標系の 1 つ目の軸は、 x_1, x_2 を通る直線に垂直な実軸 (とそれに付随する虚軸)、2 つ目の軸は、 x_1, x_2 の中点を通る直線に垂直な実軸 (とそれに付随する虚軸) である。

3. 低ランク行列の積による巡回群行列の表現

前節で行列 G の性質について述べたが、 $N = n$ を仮定したと、また $N \times N$ 行列の対角化を必要とすることが問題である。もし扱う画像が高々 128×128 画素のサイズだとしても、単純に複素行列 G を保持するだけのために膨大なメモリ領域が必要となり^(注2)、現実的ではない。

(注2): $N = 128 \times 128 = 16384$ であり、実部と虚部それぞれ double 型 (8 バ

ここでは再び $N > n$ を仮定し、 $n \times n$ の対角化での実現について議論する。

3.1 列置換行列による巡回群の表現

前節では画像列 X_0 に左から作用する行列 G を考えたが、逆に右から作用する行列を考えると、以下のように単なる $n \times n$ の列置換行列 M で表現できる。

$$X_1 = X_0 \begin{pmatrix} 0 & & & & 1 \\ 1 & & & & \\ & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & 0 \\ & & & & & & 0 \end{pmatrix} = X_0 M \quad (20)$$

すると、式 (9) は以下のように書ける。

$$G = X_0 M (X_0^T X_0)^{-T} X_0^T \quad (21)$$

ここで $\text{rank}(G) = \text{rank}(M) = n$ である。

単に画像列 X_0 の列を取り替えたいのであれば、 G よりも次元の小さい M を使ったほうが効率がよい。しかし本手法の目的は、ある一枚の画像に有効に働く行列を考えることであるため、右から作用する M をそのまま用いるのは適切ではない。

しかし、上式のように G は M を含む表現になる。そこでこれを利用して、 G の対角化の代わりに M を対角化することを考えてみる。

3.2 低ランク行列の対角化による行列 G の構成

画像列 X_0 の特異値分解を $X_0 = E \Sigma V^T$ とすると、式 (21) は

$$G = E \Sigma V^T M V \Sigma^{-1} E^T \quad (22)$$

と書き直せる。一方列置換行列 M は、複素対角行列 D_M と複素行列 W を用いて

$$M = W^{-1} D_M W \quad (23)$$

と対角化できる（詳細は[18]を参照）。これを代入し、 $U_1 = W V \Sigma^{-1} E^T$ 、 $U_2 = E \Sigma V^T W^{-1}$ とおくと、

$$G = E \Sigma V^T W^{-1} D_M W V \Sigma^{-1} E^T = U_2 D_M U_1 \quad (24)$$

となる。これは、以下のように式 (10) の固有値 0 に対応する部分を削除したものと一致する。

$$U_1 = \begin{pmatrix} I_n & 0 \end{pmatrix} U, \quad U_2 = U^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix} \quad (25)$$

$$D_M = \begin{pmatrix} I_n & 0 \end{pmatrix} D \begin{pmatrix} I_n \\ 0 \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n) \quad (26)$$

固有値 0 に対応する固有ベクトルは必要ないため、 $N \times N$ の行列 G を直接対角化する代わりに、 $N \times n$ の行列 X_0 の特異値分解と $n \times n$ の行列 M の対角化で、 G の対角化が計算できることになる^(注3)。

イト)を確保するとすると、 $16384 \times 16384 \times 8 \times 2 = 4.295 \times 10^9$ Byte = 4GB となる。また一般的に行列 G は密であり、疎行列のようなメモリ領域の削減が行えない。実際には、さらに行列のサイズに応じた計算のためのワークエリアが必要となる。

(注3): 画像列が互いに相関がある場合には、小さい特異値 Σ に対応する E の列ベクトルを削除することがよく行われるが、本手法でそれを行うことはできない（詳細は省略する）。

ここで再び G^n を考えると、式 (22), (24) と $U_1 U_2 = I_n$ から

$$G^n = U_2 D_M^n U_1 = E \Sigma V^T M^n V \Sigma^{-1} E^T \quad (27)$$

と計算できる。ここで $M^n = I_n$ であるので、

$$M^n = I_n = W^{-1} D_M^n W \quad (28)$$

$$D_M^n = W M^n W^{-1} = I_n \quad (29)$$

となる。つまり、 $N = n$ の場合と同様に、 $N > n$ の場合も D_M の 0 でない固有値はやはり 1 の原始 n 乗根になることがわかる。

したがって、前節の結論は次のように言い換えることができる。「複素行列 U_1 は、単に各画素を複素回転させることで画像が x'_1, x'_2, \dots と次々に移りあうことができる複素部分空間への投影を実現している」。

4. 新しい中間姿勢の画像の生成

4.1 複素部分空間での回転による中間画像の生成

前節で、画像列 x_1, x_2, \dots の関係を単純な 1 次元複素空間（複素平面）での回転で表現できるような部分空間への線形写像を表す複素行列 U_1, U_2 の構成方法について述べた。これにより、 x_1 から x_2 へ、また x_1 から x_3 への画像の遷移を、それぞれ $x_2 = G x_1$ と $x_3 = G^2 x_1$ という行列 G を用いた形ではなく、 $x_2 = U_2 D_M U_1 x_1$ と $x_3 = U_2 D_M^2 U_1 x_1$ という行列の積により実現することができる。

ではここで、2 乗すると G となる行列 $G^{\frac{1}{2}}$ というものを考えてみよう。式 (27) と同様にして、対角行列部分を $\frac{1}{2}$ 乗することで、そのような行列が得られる^(注4)。

$$G^{\frac{1}{2}} = U_2 D_M^{\frac{1}{2}} U_1 \quad (30)$$

$$D_M^{\frac{1}{2}} = \text{diag}(\lambda_1^{\frac{1}{2}}, \lambda_2^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}}) \quad (31)$$

この $G^{\frac{1}{2}}$ を x_1 や x_2 に作用させると、 $x_{1.5} = G^{\frac{1}{2}} x_1$ や $x_{2.5} = G^{\frac{1}{2}} x_2$ などという、サンプル間に存在するであろう中間の画像を生成することができないだろうか。

このようにして作成した画像が実際の物体を回転させたときの中間画像に一致すれば、本手法が提案する複素部分空間はまさに有用なものになるはずである。それが本研究の目的であるが、まだそこには至らないため、ここでは実際は生成される中間画像はどのようなものなのかを考察する。

図 1 を見れば明らかなように、固有値 1 に対応する軸では部分空間に投影された後でも、1 の偏角は 0 であるため、回転されていない。そのため、生成される中間画像の実部だけを見ると、元のサンプルが張る空間内に収まっている。このことは式 (24) を見ても分かるが、 G のもっとも左側に存在する E はサンプル画像が張る部分空間の基底ベクトルであり、それより右側の行列の積はその部分空間の係数（ただし複素数）となっている。したがって、中間画像の実部は元のサンプルが張る空間

(注4): 実際には $(-1)^2 = 1$ であるので、 $D_M^{\frac{1}{2}}$ の要素の表現は $\lambda_j^{\frac{1}{2}}$ と $-\lambda_j^{\frac{1}{2}}$ と 2 通りある。一般に $D_M^{\frac{1}{m}}$ の要素の表現は、1 の原始 m 乗根の数の m 通り存在する。ここではもっとも直感的な、偏角を $\frac{1}{m}$ にするものを採用している。

内に存在し、虚部はそれと直交する空間に飛び出るようにして回転していることを意味している。

4.2 画像の生成例

これまでに述べた手法を用いて、実際の画像列における複素部分空間を構成し、中間画像を生成した結果を示す。画像は COIL-20 [22] の物体 4 (招き猫) を用いた。この画像のサイズは $N = 128 \times 128$ であり、5 度ずつ 1 軸回転した $n = 72$ 枚の画像サンプルからなる。手法の実装と複素対角化の計算には Scilab-4.1 を用いた。

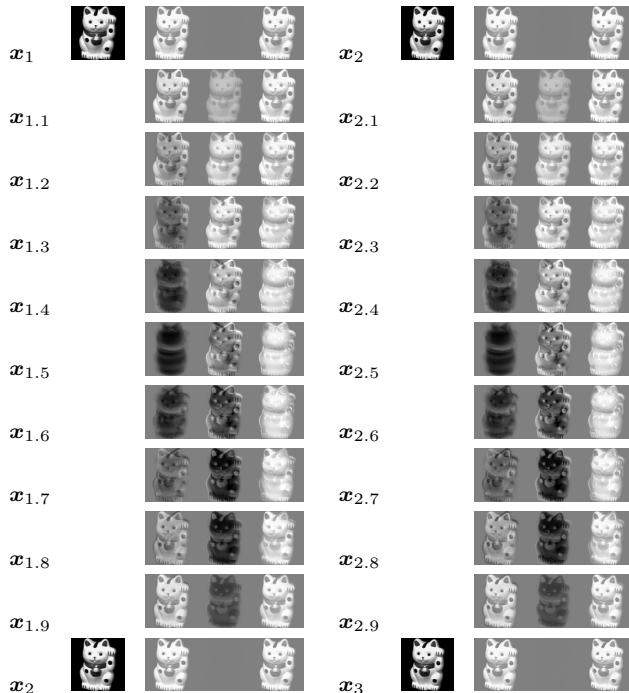


図 2 行列 $G^{\frac{1}{10}}$ を画像へ作用させて作成した補間画像。左から 1 列目は原画像で、coil-20(obj4) の 0, 5, 10 度の画像 x_1, x_2, x_3 のみを示す。2 列目から 4 列目は、それらの間を $\frac{1}{10}$ にして補間した複素画像 $x_{1+0.1j} = G^{\frac{j}{10}} x_1$ ($j = 0, 1, \dots, 20$) の実部、虚部、絶対値を示す。灰色が 0 で、正の値を白く、負の値を暗く表示している。

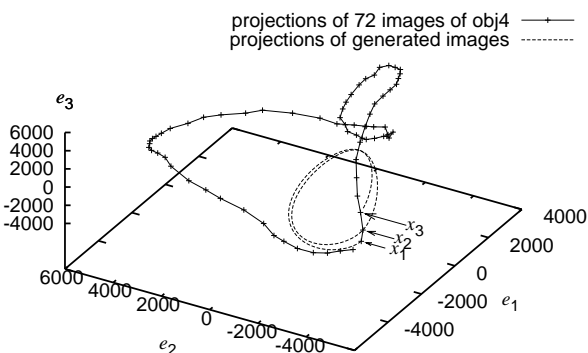


図 3 固有空間における coil-20(obj4) の 72 枚の原画像と生成した補間画像の軌跡。補間画像については、coil-20(obj4) の 0, 5, 10 度の画像 x_1, x_2, x_3 の間のみの補間を示す。

72 枚の画像から複素部分空間を計算し、生成した中間画像の一部を図 2 に示す。学習サンプル画像のうち 0, 5, 10 度に対応する画像 x_1, x_2, x_3 の間を 10 等分したときの中間画像 $x_{1.1}, x_{1.2}, \dots$ を計算した。これは複素画像になるため、それぞれの画像の各画素の実部と虚部と絶対値を画像として表示している。学習サンプルと一致する部分では虚部は 0 となり、それ以外の中間画像では実部と虚部が徐々に正負を交替するように (白っぽい画像から暗い画像へと) 入れ替わっているのがわかる。とくに実部画像はサンプル間の中間点において平均画像に近いものになっており、絶対値画像ではサンプル間がぼけて補間されているようなものになってしまっている。

このことを確かめるため、中間画像を画像列の固有空間に投影した結果を図 3 に示す。固有空間は 72 枚の画像から生成したもので、各軸の e_1, e_2, e_3 は第 1, 2, 3 固有ベクトルを表す。72 枚の画像の投影点列を結んだものを実線で示しており、 x_1, x_2, x_3 は矢印で示してある。そして x_1, x_2, x_3 の間を 100 等分して生成した中間画像の絶対値を投影した軌跡を点線で示してある。この軌跡は x_1 から出発し、大きく円を描くようにして平均画像を表す原点付近を通過し、 x_2 へと到達している。この図では表示はしていないが、すべてのサンプル間の中間画像を投影した結果、やはり同様に螺旋を描くような軌跡となっている。このように、サンプルの間ではサンプル画像から離れて原点へと向かう軌跡になってしまっているために、図 2 のように画像としてみた場合に、ぼやけた平均画像のような中間画像を経由してサンプル間を遷移してしまっている。

本研究の目的である、姿勢推定や補間に有用な部分空間の構成という観点から見れば、この結果は明らかに満足できるものではない。もっとサンプル間を直接的に結ぶような軌跡を生成する部分空間を構成するために、今後さらなる検討が必要である。

一方で、画素数がサンプル数より多い場合の線形手法を追求するという点においては、本手法の有用性が見出せる。すなわち、これまでは画像列が強い非線形性を持っているために、多様体学習や非線形手法が必要だと思われていたが、本手法では線形手法によって学習サンプル間を滑らかに補間する軌跡を生成することができている。そのため、今後本手法を改良して直感に沿ったサンプル画像の補間を行う部分空間が構成できれば、線形手法による画像の姿勢推定が行えるのではないかと考えている。

5. おわりに

本論文では、1 軸回転する物体のような巡回する画像間の関係を単純な複素空間での回転で表現できるような複素部分空間への線形写像を提案し、低ランク行列の積によるその線形写像行列の構成方法と、画像補間への応用を示した。補間画像の生成例で示したように、まだまだ本手法による部分空間が有用なものであるとは言えず、検討すべきことが山積している。

まず一つ目は、図 3 で示したように、中間画像の軌跡が平均画像付近を経由してしまっている問題の検討である。これは、共役複素数が対で固有値になっていることを考慮していない

めであり、実ブロック対角化手法 [17], [18] のようにそれらに対して用いるべきである。また、行列 G を構成する際に、式 (9) の最小ノルム型の一般化逆行列を用いていることも一因とみられる。この一般化逆行列は $N - n$ 次元の零空間を持っているため、サンプル画像の情報を用いて零空間の中からより適切な解を探索するという手法 [23] を応用することを考えている。

また一般化逆行列を用いずに、そもそも式 (17), (18) のように「投影したら複素回転で画像列の遷移が表現できる複素部分空間が存在する」という仮定から出発して、 U_1 を構成する n 個の行ベクトル $u_k (k = 1, \dots, n)$ を以下のように推定することも考えられる。

$$\min \sum_j |x'_{j+1k} - \lambda_k x'_{jk}|^2 = \sum_j |u_k x_{j+1} - \lambda_k u_k x_j|^2 \quad (32)$$

こうすることで、部分空間への U_1 の推定に、画像間の遷移の滑らかさなどの知識を制約条件として入れることができると思われる。

二つ目は、姿勢推定への応用である。既知の物体の新たな画像 x が与えられたときに、すでに学習しているサンプル画像 x_1 とどれだけ姿勢が変わっているかということがこの部分空間から分かると思われる。つまり x と x_1 を部分空間に投影し、その偏角の差が行列 G の何乗で表せるか、つまり $U_1 x = D_M^\alpha U_1 x_1$ を満たす適切な α を推定する問題に置き換えることができる。

三つ目は、3 次元の回転の推定問題への拡張である。単純な巡回群として学習サンプルを扱うだけでは、1 軸回転以上の物体の姿勢を扱うことはできない。そこで 3 次元の姿勢の回転を表すための四元数を用いて、先ほどと同様に「投影したら四元数による回転で画像列の遷移が表現できる四元数体の部分空間が存在する」という仮定から出発することで、部分空間への線形写像を推定できるのではないかと考えている。また四元数体の行列の対角化 [24] ~ [26] も研究されており、3 次元回転を表す行列の対角化も検討の余地があると思われる。

以上のような問題点はあるものの、本手法は線形手法を再考するきっかけになりうるものであると思われる。次元数がサンプル数よりもかなり大きいという条件は、認識問題においては次元の呪いとして知られているが、非線形手法はその条件を利用して高次元空間における線形判別を実現している。本手法が提案するように、その条件は姿勢推定においても有効に利用することができるのではないかとと思われる。また、これまで画像認識において用いられることのなかった複素部分空間の積極的な利用が問題を解決する鍵になりうる、ということ示唆に富んだ考え方であると考えている。

文 献

[1] Lowe, D. G.: Fitting Parameterized Three-Dimensional Models to Images, *IEEE Trans. PAMI*, Vol. 13, No. 5, pp. 441-450 (1991), online.
 [2] 村瀬洋, シュリーナイヤー: 2 次元照合による 3 次元物体認識 パラメトリック固有空間法, 電子情報通信学会論文誌 DII, Vol. J77-D2, No. 11, pp. 2179-2187 (1994), online.
 [3] Murase, H. and Nayar, S. K.: Visual learning and recognition of 3-D objects from appearance, *Intl. J. of Computer Vision*, Vol. 14, No. 1, pp. 5-24 (1995), online.

[4] Scholkopf, B., Smola, A. and Müller, K.-R.: Nonlinear Component Analysis as a Kernel Eigenvalue Problem, *Neural Computation*, Vol. 10, pp. 1299-1319 (1998), online.
 [5] 坂野鋭一: パターン認識における主成分分析 顔画像認識を例として, 統計数理, Vol. 49, No. 1, pp. 23-42 (2001), online.
 [6] Pontil, M. and Verri, A.: Support Vector Machines for 3D Object Recognition, *IEEE Trans. PAMI*, Vol. 20, No. 6, pp. 637-646 (1998), online.
 [7] Law, M.: Manifold Learning, online (accessed 2007.3.20), online.
 [8] Turk, M. A. and Pentland, A. P.: Face recognition using eigenfaces, *CVPR91*, pp. 586-591 (1991), online.
 [9] Pentland, A., Moghaddam, B. and Starner, T.: View-based and modulear eigenspaces for face recognitoin, *CVPR94*, pp. 84-91 (1994), online.
 [10] Smola, A. J. and Schölkopf, B.: A Tutorial on Support Vector Regression, *NeuroCOLT Technical Report NC-TR-98-030*, Royal Holloway College, University of London, UK (1998), online.
 [11] Cristianini, N. and Shawe-Taylor, J.: サポートベクターマシン入門, 共立出版 (2005), online.
 [12] 安藤慎吾, 草地良規, 鈴木章, 荒川賢一: サポートベクトル回帰を用いた三次元物体の姿勢推定法, 電子情報通信学会論文誌 D, Vol. J89-D, No. 8, pp. 1840-1847 (2006), online.
 [13] Okatani, T. and Deguchi, K.: Yet Another Appearance-Based Method for Pose Estimation Based on a Linear Model, *IAPR Workshop on Machine Vision Applications 2000*, pp. 258-261 (2000).
 [14] 天野敏之, 玉木徹: Estimation-by-Completion: 3 次元物体の線形姿勢推定手法, *MIRU2006 画像の認識・理解シンポジウム論文集*, pp. 460-465 (2006), online.
 [15] 出口光一郎, 岡谷貴之: 固有空間法はなぜうまく働くか, 情報処理学会コンピュータビジョンとイメージメディア研究会, Vol. 2001, No. 66, pp. 1-8 (2001), online.
 [16] 玉木徹, 天野敏之: マルチポート固有空間法, 部分空間法研究会 2006, pp. 7-15 (2006), online.
 [17] 玉木徹, 天野敏之, 金田和文: 巡回群行列のブロック対角化を用いた回転物体画像の部分空間への投影とその応用, *MIRU2007 画像の認識・理解シンポジウム論文集* (2007), online.
 [18] 玉木徹, 天野敏之, 金田和文: 画像列の遷移行列のブロック対角化による部分空間を用いた物体の姿勢推定, *MIRU2007 画像の認識・理解シンポジウム論文集* (2007), online.
 [19] 志賀浩二: 群論への 30 講, 朝倉書店 (1989).
 [20] Menke, W.: 離散インバース理論 — 逆問題とデータ解析, 古今書院 (1997), 柳谷 俊, 塚田 和彦 訳.
 [21] 志賀浩二: 固有値問題 30 講, 朝倉書店 (1991).
 [22] Nene, S. A., Nayar, S. K. and Murase, H.: Columbia Object Image Library (COIL-20), Technical Report CUCS-005-96, Columbia University (1996), online.
 [23] 篠村祐司, 玉木徹, 天野敏之, 金田和文: 固有空間による人物の歩行軌跡の予測と零空間を用いた予測修正, 電子情報通信学会パターン認識・メディア理解研究会, Vol. PRMU2006-239, pp. 25-30 (2007), online.
 [24] Bunse-Gerstner, A., Byers, R. and Mehrmann, V.: Numerical methods for simultaneous diagonalization, *SIAM Journal on Matrix Analysis and Applications*, Vol. 14, No. 4, pp. 927-949 (1993), online.
 [25] Tongsong, J. and Li, C.: Generalized diagonalization of matrices over quaternion field, *Applied Mathematics and Mechanics*, Vol. 20, No. 11, pp. 1297-1304 (1999), online.
 [26] Jiang, T.: Algebraic methods for diagonalization of a quaternion matrix in quaternionic quantum theory, *Journal of Mathematical Physics*, Vol. 46, No. 5, p. 052106 (2005), online.